The Role of Trusted Relationships on Content Spread in Distributed Online Social Networks

Valerio Arnaboldi, Massimiliano La Gala, Andrea Passarella, and Marco Conti

IIT-CNR, Via G. Moruzzi 1, 56017, Pisa, Italy
{v.arnaboldi,m.lagala,a.passarella,m.conti}@iit.cnr.it

Abstract. In Distributed Online Social Networks (DOSN) content spread will largely depend upon trust relationships between users, who are likely to allocate resources only to help spreading content coming from peers with whom they have a strong enough relationship. This could lead to the formation of isolated groups of intimates in the network, and to the lack of a big enough connected component, essential for the diffusion of information. In this paper we simulate the outcome of such restrictions by using a large-scale Facebook data set, from which we estimate the trust level between friends. We then simulate content spread on the same network assuming that no central control exists, and that social friendship links exist only above certain levels of trust. The results show that limiting the network to "active social contacts" of the users leads to a high node coverage. On the other hand, the coverage drops for more restrictive assumptions. Nevertheless, selecting a single excluded social link for each user and adding the respective node in the network is sufficient to obtain good coverage (i.e. always higher than 40%) also in case of strong restrictions.

Keywords: Distributed online social networks, Trust based communications, Information diffusion.

1 Introduction

Online Social Networks (hereinafter OSN) like Facebook and Twitter are becoming essential everyday tools useful in many different situations, from the management of personal social relationships to advertisement and professional networking. This makes them one of the most important cases nowadays of large-scale virtual environments. Despite this, the fast pace at which OSN are growing rises some fundamental questions regarding the sustainability of their architectures. In fact, OSN are generally based on centralised solutions, that guarantee more control upon user's data and consequently generate more value for the service providers. Nevertheless, the large amount of communication data generated by OSN requires huge storage capacity and complex solutions for providing instant access to the users. In addition, from a more ethical point of view, OSN often use the personal data of their users for commercial purposes, centralising, together with the data, also wealth, possibly resulting in power law economies [12].

L. Lopes et al. (Eds.): Euro-Par 2014 Workshops, Part I, LNCS 8805, pp. 287–298, 2014.

Several distributed alternatives to OSN have been proposed in the last years, broadly identified by the term Distributed Online Social Networks (DOSN). DOSN (e.g. diaspora* or PeerSon [6]) replicate OSN features in a fully decentralised way. Specifically, DOSN permit to manage a *digital personal space* for each user, where the latter can leave or receive asynchronous messages or post other kind of personal information. Moreover, DOSN support the creation of *social links* between users, giving different access policies to digital personal spaces for friends compared to strangers. DOSN also provide instant messaging functionality in the form of private communications.

In DOSN, each user maintains her personal data locally or on intermediate servers, and interactions between users occur through peer-to-peer (P2P) communications. Compared to OSN, DOSN provide the user with much more control over her personal data, and data decentralisation guarantees privacy and low complexity. Since in DOSN data are completely decentralised and rely upon the personal devices of the users the circulation of information cannot be controlled via a centralised server. For this reason, content can only be disseminated in the network through chains of social links between users. Clearly, users are inclined to favour the dissemination of content coming from trusted social contacts (i.e. with a strong enough relationship), which generally share similar interests with the users due to the effect of homophily, for which people tend to bond with similar alters [7]. In addition, the restriction of information diffusion to trusted peers lowers the amount of content coming from all possible social links without discrimination that could waste local resources contributed by users' devices to support DOSN functionality. Therefore, it is reasonable to assume that DOSN users will be willing to help replicating and disseminating content coming only from a set of users they trust most, and discarding the rest. This is clearly a double-edge sword, because it also limits the spread of information to possible interested users, and may potentially reduce data availability. It is therefore important to understand the effect of this type of filtering on content dissemination in DOSN.

In this paper we simulate the impact of the restriction of communication to trusted contacts in DOSN on information diffusion in the network. To do this, we study the topology of the graph induced by the restriction of social links to trusted relationships only. In particular, we look at the size and the number of components containing connected nodes in the graph. Each component represents a portion of the original network through which information can reach all the connected nodes. Clearly, different network components (and thus different information dissemination patterns) emerge depending on whether more or less strict trust restrictions are considered. Since the collection of network graphs representing DOSN is difficult, for the distributed nature of the system, we take a larger scale Facebook graph as starting point. We estimate the trust level between connected users through the frequency of interactions between them (which is well backed-up by results in sociology [11,4]), and simulate content diffusion on this network considering different thresholds for defining trusted links. We assume that no central control exists on this network and we select the

set of trusted contacts for its users according to their contact frequency, then studying the properties of the resulting graph. Specifically, if a social relationship is not trusted (i.e. it does not have a sufficiently high contact frequency), the respective social link is not included in the graph we used in simulation. To assess the impact of the selection of trusted links, we define the minimum level of trust by setting a threshold on the contact frequency of the links to be included in the graph. We take values of this threshold equal to the frequencies of contact that have been used in the literature for defining different levels of social relationships [11]. In particular, we consider the well known ego network model [14], whereby social relationships of a user (ego) can be divided in concentric layers of increasing size and decreasing social intimacy (i.e. corresponding to decreasing tie strength and fewer interactions) In this way we obtain different social graphs with different minimum levels of trust, that coincide with a natural categorisation of social relationships in humans. Note that this way of estimating trust lends itself to automatic systems to decide on which social links to accept content, just by monitoring the frequency of interactions on them.

The results of the analysis indicate that limiting content spread to social contacts that coincide with the definition of "active social contacts" of the users, which corresponds to the most external layer in the ego network model, leads to a network graph with a sufficiently large component of connected nodes, which covers more than 96% of the original Facebook network. Restricting content spread to the next layer of the ego networks, or further, makes the relative size of the biggest connected component (and therefore coverage) drop below 30%. Since the remaining components are very small compared to the largest one for all the used thresholds, diffusing information in the network could be problematic when the largest component does not cover a sufficiently high number of nodes. As a possible solution to increase node coverage in case of very restrictive thresholds we investigate the effect of adding to the graph only one social contact for each user, selected with different possible strategies (e.g. select the link with highest/lowest contact frequency with the user, select a random acquaintance, etc.). The results indicate that this solution considerably increases the number of covered nodes, even in case of very strong trust. Noticeably, the best strategy is a probabilistic selection of a social contact based on their contact frequency with the considered user, whereas taking always the contact with highest/lowest contact frequency (below the minimum contact frequency imposed by the restriction) leads to worse results in terms of number of nodes covered. Clearly, adding a contact to the list of trusted nodes represents a cost for the users in terms of additional unwanted content, but limiting the choice to a single node should be a reasonable solution for them since they would receive a global return in terms of quality and quantity of information circulating in the network.

The paper is organised as follows: in Sect. 2 we describe the Facebook data set we have used for the analysis. Hence, in Sect. 3 we introduce the methodology we use, in particular how we define the values of the threshold we use in the analysis. Then, in Sect. 4 we report the results of the analysis. In Sect. 5 we

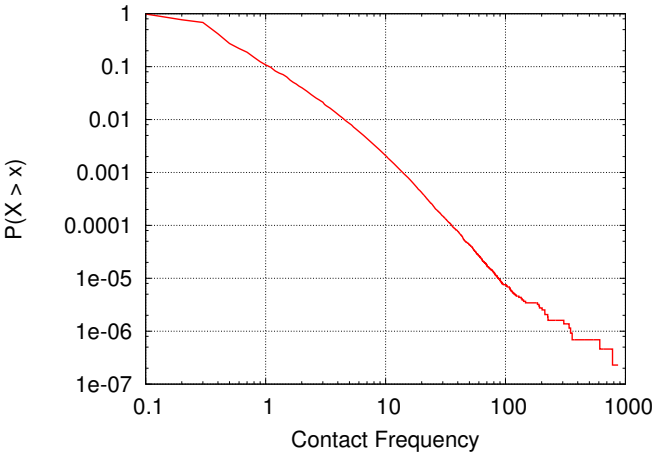

Fig. 1. CCDF of the contact frequency for the links

describe the related work in the literature, especially the most relevant DOSN solutions already existing. Last, in Sect. 6 we draw the conclusions of our work.

2 Data Set Description

To perform our analysis, we use a large-scale Facebook data set containing information about social interactions between users. The data set represents a large Facebook regional network, with more than 3 million users and more than 23 million social links between them, which has been downloaded in 2008, when the default privacy policy allowed users within a regional network (a feature removed by Facebook in 2009) to see all the personal data of other users in the same network. The data set is publicly available for research[1] and it has been largely used for other analyses since it shows typical social network characteristics [15,3,1].

The data set contains information about the existence of social links between users, that are represented by a *social graph*, where nodes represent Facebook users and links Facebook friendships between them. On the other hand, social interactions recorded in the data set are represented by a series of *interaction graphs*, each of which contains the number of interactions occurred on the social links within temporal windows of increasing duration. The interaction graphs identify communications occurred respectively *one month*, *six months*, *one year* before the download, and for the *all* the duration of the relationships. By combining these graphs, we have been able to estimate the contact frequency between users, following the procedure described in [3], obtaining a unique graph where the links are weighted by the contact frequency between the involved users.

[1] `http://current.cs.ucsb.edu/socialnets/`

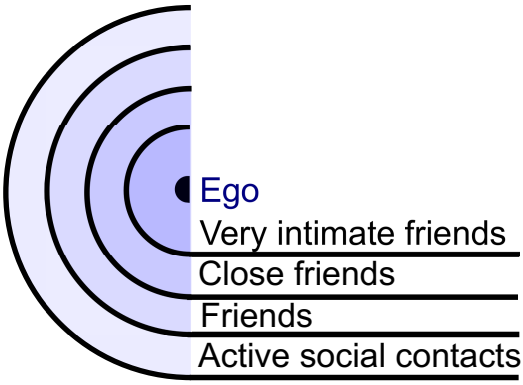

Fig. 2. Ego network model

In this work we use the contact frequency between users in Facebook as a proxy for the level of trust between them. This is supported by results in the literature that identified a strong relation between the contact frequency and the tie strength or emotional closeness between people, both in offline and online environments [4,11,13]. The complementary cumulative distribution function (CCDF) of the contact frequency for the links in the graph obtained from the data set is depicted in Fig. 1. The figure indicates that the distribution has a power law trend, thus implying that most of the links in the network have a very low level of trust, whereas only few links have very high trust. For this reason, we expect that restricting the network to trusted links only could have a strong impact upon the structural properties of the resulting graph.

3 Social Networks for Content Diffusion

Since in our analysis we are interested in users who actively communicate with others, we select from the Facebook graph only the users with at least one active link (i.e. with contact frequency > 0) and we discard all the other users, that indeed are inactive. Moreover, we further restrict the analysis to the set of users that have communicated with other users at least 6 months before the time the data set was downloaded. This ensures that our analysis is restricted to sufficiently stable users. In fact, the contact frequency of new users in OSN is generally higher that that of older (and more stable) users [2] and could bias the analysis. The resulting graph, after this pre-process, consists of $1,083,209$ nodes and $7,709,309$ links.

To simulate the restriction of communication to a list of trusted contacts for each user in DOSN we apply a series of filters to the Facebook social graph previously described, eliminating the links with contact frequency below the chosen threshold, that defines the boundary of the trusted contact list.

3.1 Trusted Contact List Based on Contact Frequency

In the literature, social relationships of a person are arranged in a series of inclusive concentric layers around her, with sensibly different properties. These layers are defined by the frequency of contacts between users, and therefore we can assume that the layered structure also represents the structure of trust of social relationships. Specifically, the ego network model (depicted in Fig. 2) defines four circles of alters around the ego (i.e. the considered individual) [14]. The first and innermost layer is the *support clique*, containing on average five people very close to the ego and contacted by her at least once a week. People in this layer can be broadly identified as "very intimate friends". The *sympathy group* (that includes the support clique) contains fifteen members contacted at least once a month, which can be identified as "close friends". The *affinity group* contains fifty members contacted at least $\sim$ eight times a year [3]. People in this layer can be defined as "friends". Lastly, the *active network* contains 150 people contacted at least once a year. These people are those for whom the ego invests a non negligible amount of cognitive resources and can be defined as "active social contacts". Beyond the active network, alters are mere acquaintances, and their social relationships are not actively maintained by the ego.

Based on the definition of the ego network layers we can identify possible trusted contacts lists definitions that could be adopted in DOSN. For example, to simulate the presence of lists containing "friends" we can fix the minimum contact frequency to be considered in the analysis to eight messages a year. In the analysis we use all the values of contact frequency that define the ego network layers previously presented to identify four different possible definitions of trusted contacts list and we discuss the implications of using these filtering strategies in DOSN. Specifically, in the rest of the paper we indicate the values of the thresholds in number of messages per month, so "1/12" represents one message a year, "8/12" eight messages a year, "1" one message per month, and "4" four messages per month.

3.2 Network Connectivity

Having defined the values of the threshold to generate the network graphs at different levels of trust, we proceed the analysis of connectivity of these graphs to find out if they are suited for information diffusion. Specifically, for each graph we study the size of the largest component of connected nodes and we compare it to the number of nodes in the original graph (the one obtained after the pre-processing phase). In addition, we study the distribution of the size of the remaining components of connected nodes (not considering the largest one). This is useful for identifying the number of components that must be reached by information to cover a sufficiently large portion of the network.

As will be clear from the results in Sect. 4, for some threshold on the trust level we obtain quite small largest connected components, and a big number of extremely small additional disconnected components. To improve content spreadability of the network we tried, as a possible alternative to lowering the trust

Table 1. Percentage of nodes of the original graph covered by the largest component for the different thresholds. Thresholds are expressed in msg/month.

Threshold	Percentage of nodes in the largest component					
	No insert	**High freq**	**Low freq**	**Prob**	**Inv. prob**	**Rand**
1/12	0.966	0.994	0.994	0.994	0.994	0.994
8/12	0.297	0.714	0.705	0.726	0.722	0.725
1	0.191	0.642	0.634	0.661	0.657	0.661
4	0.028	0.386	0.385	0.453	0.444	0.456

value of the system, the re-insertion of one social contact for each user in the graph obtained at each level of trust. We tested several possible strategies to select this social contact for each user. Specifically, the strategies that we use are the following: (i) selection of the contact with the highest contact frequency; (ii) selection of the contact with lowest contact frequency; (iii) extraction with probability proportional to the contact frequency; (iv) extraction with probability inversely proportional to the contact frequency; (v) random uniform extraction. We assess the impact of the re-insertion of the social contacts on the global connectivity of the resulting graph (i.e. the largest component of the graph) and we evaluate the pros and cons of the adoption of the different strategies.

4 Results

The proportion of nodes of the original graph (i.e. the one obtained after the preprocessing phase described in Sect. 3) covered by the largest component for each threshold is reported in Tab. 1 in the column "No insert", which indicates that we have not applied any re-insertion strategy on these results. The first largest component obtained with the threshold coinciding with the contact frequency of "active social contacts" (as defined in Sect. 3.1) guarantees a very high node coverage, particularly favourable condition for information diffusion. With the second threshold, related to "friends", the number of nodes covered by the largest component drops to $\sim 30\%$. This could still be enough to diffuse information to a sufficiently high number of nodes. Nevertheless, for the other values of the threshold the node coverage is too low.

If we look at the distribution depicted in Fig. 3, representing the size of the components of connected nodes in the network excluded the largest one, we can notice that for all the thresholds, and especially in case of no re-insertion, these components are always really small, especially when compared to the largest one. In fact, the distributions show power-law trends with a maximum component size of 95 nodes. This indicates that if we want to reach a high number of nodes in the network and the largest component is not sufficient to do so, it is necessary to place information on a big number of additional components, without relying on automatic spread of information over trusted social links. This is further confirmed by the results in Tab. 2, 3, 4 and 5 under the row "No insert", that indicates the number of components that must be infected by

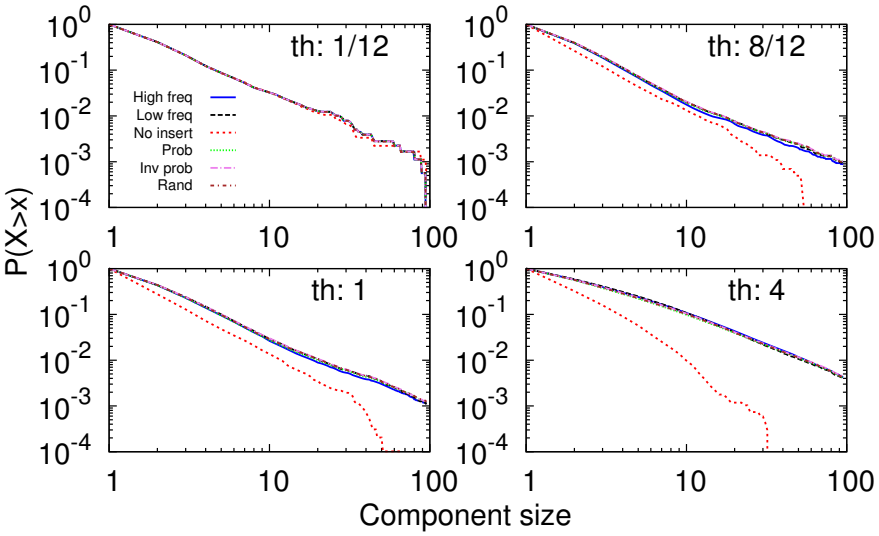

Fig. 3. CCDF of the size of the components for each threshold and strategy, excluded the largest component

information to reach the desired level of coverage, expressed in percentage with respect to the number of nodes in the original network. Also from these tables a sufficient coverage is obtained only with the first threshold (one message a year in Tab. 2), whereas for the other thresholds the number of components needed to reach a reasonable coverage (e.g. above 50% of the network) is very high and would result in a very expensive process. Moreover, whilst it could be relatively easy to identify the largest component in the network, it is not easy to identify all the remaining components, especially in decentralised systems like DOSN. This fact could further limit the diffusion process.

In case the level of trust needed in DOSN were too restrictive to reach a sufficient coverage for the dissemination of information, We evaluate a possible alternative to the mere decrease of the trust level, that would result in a less

Table 2. Number of components needed to cover the specified percentage of nodes in the original network using a threshold equal to 1/12 msg/month

	Coverage						
Strategy	40%	50%	60%	70%	80%	90%	100%
No insert	1	1	1	1	1	1	31,987
High freq	1	1	1	1	1	1	1,784
Low freq	1	1	1	1	1	1	1,784
Prob	1	1	1	1	1	1	1,784
Inv prob	1	1	1	1	1	1	1,784
Rand	1	1	1	1	1	1	1,784

Table 3. Number of components needed to cover the specified percentage of nodes in the original network using a threshold equal to 8/12 msg/month

	Coverage						
Strategy	40%	50%	60%	70%	80%	90%	100%
No insert	94, 218	202, 539	310, 860	419, 181	527, 502	635, 823	744, 143
High freq	1	1	1	1	43, 045	151.366	259, 686
Low freq	1	1	1	1	61, 369	169, 690	278, 010
Prob	1	1	1	1	37, 623	145, 944	254, 264
Inv prob	1	1	1	1	45, 197	153, 518	261, 838
Rand	1	1	1	1	40, 343	148, 664	256, 984

trusted system. Specifically, we add, for each user, a single additional social contact to the list of contacts trusted by the user. We apply the latter solution on the Facebook graph at each threshold, testing the different re-insertion strategies described in Sect. 3. In Tab. 1, from the third to the last column, we report the size of the largest component of connected nodes for each combination of threshold and re-insertion strategy. As can be noted, the impact of the re-insertion is substantial for thresholds $> 1/12$. For the threshold of 1/12 the re-insertion is not needed, since most of the nodes of the original network are already present in the resulting graph. Even for the most restrictive threshold (4 messages per month) the gain due to the re-insertion brings the node coverage to $\sim 40\%$.

The results of the different strategies vary significantly, with the probabilistic and the random strategies ("Prob" and "Rand" in the tables) giving the highest improvement in terms of number of nodes covered, as reported in Tab. 1. In addition, as reported in Tab. 2, 3, 4 and 5, these two strategies are the most convenient also when all the other components, in addition to the largest one, are considered. This result is perhaps not surprising since the two fixed strategies related to the selection of the node with highest/lowest contact frequency with the considered user ("High freq" and "Low freq" in the table) often lead to a node too close to the user and with too many social contacts in common, that do not help to improve the number of covered nodes, or that, on average, has very low contact frequency also with other nodes, and thus introduces few

Table 4. Number of components needed to cover the specified percentage of nodes in the original network using a threshold equal to 1 msg/month

	Coverage						
Strategy	40%	50%	60%	70%	80%	90%	100%
No insert	208, 769	317, 090	425, 411	533, 732	642, 053	750, 374	858, 694
High freq	1	1	1	8, 801	87, 540	195, 861	304, 181
Low freq	1	1	1	13, 147	106, 719	215, 040	323, 360
Prob	1	1	1	4, 271	79, 561	187, 882	296, 202
Inv prob	1	1	1	5, 470	87, 379	195, 700	304, 020
Rand	1	1	1	4, 343	81, 852	190, 173	298, 493

Table 5. Number of components needed to cover the specified percentage of nodes in the original network using a threshold equal to 4 msg/month

	Coverage						
Strategy	40%	50%	60%	70%	80%	90%	100%
No insert	391, 174	499, 495	607, 816	716, 137	824, 458	932, 779	1, 041, 099
High freq	45	2, 332	12, 660	47, 708	144, 729	253, 050	361, 370
Low freq	68	3, 022	15, 938	59, 169	167, 490	275, 811	384, 131
Prob	1	431	6, 717	36, 881	132, 106	240, 426	348, 746
Inv prob	1	608	7, 708	40, 266	140, 250	248, 571	356, 891
Rand	1	396	6, 538	36, 785	133, 350	241, 672	349, 992

nodes to the giant component. The inverse probabilistic strategy ("Inv prob" in the tables) favours contacts with too low contact frequency and, for the same reason of the "Low freq" strategy, introduces fewer other nodes connected to the largest component in the network. Therefore, the probabilistic and the uniform random choice seem the best strategies to adopt. However, the uniform random strategy, although being the easiest to implement, could lead more frequently to the selection of complete strangers to the users, and this could be a non desirable solution, as it maximises the probability for users to receive irrelevant content, due to the homophily argument. The probabilistic strategy prefers those contacts that are close to the user. Consequently, this strategy seems more reasonable compared to a pure uniform random choice.

We finally look at how the different re-insertion strategies impact on the distributions of the sizes of network components other than the largest one (see Fig. 3). All strategies, for each threshold, produce a similar distribution of the size of the components, excluded the largest one. Nevertheless, the distributions vary from the case in which no re-insertion is applied, especially for restrictive thresholds (1 message per month and 4 messages per month). This can be explained by the fact that for these thresholds the largest component is sensibly smaller than for the other thresholds, and the probability of re-inserting a node connected to this component is lower. Thus, there is the presence of a higher number of larger components disconnected from the largest one.

5 Related Work

DOSN were born in recent years to address privacy concerns over OSN. Diaspora*[2] is probably the most famous DOSN nowadays. Diaspora* supports the possibility of either creating a server (called *pod*) where the user can host her personal data or using an already existing one. Social interactions are carried out through a P2P system that makes users communicate directly with each other, without passing through a single centralised server. The authors of [6] propose a similar solution, that has been also extended to be used in case of absence of

[2] `https://diasporafoundation.org/`

stable Internet connectivity [5], a scenario particularly suited for mobile devices. In [9], the authors propose a DOSN based on the automatic identification, for each user, of her ego network layers (previously defined in 3.1), using the contact frequency between the user and her social contacts. The differences in terms of trust between the different layers are used to automatically adjust the privacy policies towards the people in the layers. Moreover, the personal social network of each user is limited to her "active network", and people beyond it are excluded from the main features of the system. The solutions proposed in [10] and [8] further exploit trust relationships arranged in concentric layers around the users to replicate the data of the user on her friend's devices, guaranteeing the access to her data even though her device were inaccessible due to a temporary disconnection or turnoffs.

6 Conclusions

In this paper we investigate the possible impact of the restriction of communications in DOSN to trusted social relationships only. This restriction is essential in DOSN, since the users are willing to distribute information coming only from trusted peers to limit the resources they dedicate to communications, that are generally limited. To perform the analysis we study the topological properties of the social graph generated by DOSN with such restrictions, looking for the presence of a large component of connected nodes (at a certain threshold of trust), within which information can spread and possibly reach all its nodes. On the other hand, disconnected small components and isolated nodes represent portions of the network that are difficult to reach and that will limit the diffusion of information.

Since the collection of a social graph representing DOSN is not easy, for the distributed nature of the system, we perform our analysis on a large-scale Facebook graph and, assuming that no central control exists upon it, we limit it selecting only links above a certain level of trust, estimated through the contact frequency between users. Hence, by applying four different thresholds, used in the literature as a natural classification for human social relationships, as the minimum level of trust in the network, we study the connectivity of the resulting graph. The results indicate that for the threshold representing "active social contacts" for the users, the resulting graph is highly connected and contains a large component covering more than 96% of the original network. On the other hand, for more restrictive thresholds, the node coverage drops significantly.

To overcome a situation, where diffusing information could be problematic due to the selected threshold on trust, we propose, in addition to a mere reduction of the global level of trust in the system, a different approach based on the addition to the list of trusted nodes of each user a single social contact, that was outside the list. We investigate different strategies to select this social contact, from a pure uniform random selection to the selection of the one with highest/lowest contact frequency with the considered user. The most effective strategies are a probabilistic selection of the contact based on its contact frequency and a

uniform random selection. Nevertheless, the former is preferable since the latter could add more people that are unknown to the user, a non desirable situation.

Acknowledgements. This work is supported by the European Commission under the EINS (FP7-FIRE 288021) project.

References

1. Arnaboldi, V., Conti, M., La Gala, M., Passarella, A., Pezzoni, F.: Information Diffusion in OSNs: the Impact of Nodes' Sociality. In: SAC 2014, pp. 1–6 (2014)
2. Arnaboldi, V., Conti, M., Passarella, A., Dunbar, R.I.M.: Dynamics of Personal Social Relationships in Online Social Networks: a Study on Twitter. In: COSN 2013, pp. 15–26 (2013)
3. Arnaboldi, V., Conti, M., Passarella, A., Pezzoni, F.: Analysis of Ego Network Structure in Online Social Networks. In: SocialCom 2012, pp. 31–40 (2012)
4. Arnaboldi, V., Guazzini, A., Passarella, A.: Egocentric Online Social Networks: Analysis of Key Features and Prediction of Tie Strength in Facebook. Computer Communications 36(10-11), 1130–1144 (2013)
5. Buchegger, S.: Delay-Tolerant Social Networking. In: Extreme Workshop on Communication, pp. 1–2 (2009)
6. Buchegger, S., Schioberg, D., Vu, L.H., Datta, A.: PeerSoN: P2P Social Networking Early Experiences and Insights. In: SocialNets, pp. 46–52 (2009)
7. Curry, O., Dunbar, R.I.M.: Do birds of a feather flock together? The relationship between similarity and altruism in social networks.. Human Nature 24(3), 336–347 (2013)
8. Cutillo, L.A., Molva, R., Strufe, T.: Safebook: A Privacy-Preserving Online Social Network Leveraging on Real-Life Trust. IEEE Communications Magazine, 94–101 (December 2009)
9. Guidi, B., Conti, M., Ricci, L.: P2P architectures for distributed online social networks. In: International Conference on High Performance Computing & Simulation (HPCS), pp. 678–681 (2013)
10. Han, L., Nath, B., Iftode, L., Muthukrishnan, S.: Social Butterfly: Social Caches for Distributed Social Networks. In: SocialCom 2011, pp. 81–86 (2011)
11. Hill, R.A., Dunbar, R.I.M.: Social network size in humans. Human Nature 14(1), 53–72 (2003)
12. Lanier, J.: Who Owns the Future? (2013)
13. Marsden, P.V., Campbell, K.E.: Measuring Tie Strength. Social Forces 63(2), 482–501 (1984)
14. Sutcliffe, A., Dunbar, R., Binder, J., Arrow, H.: Relationships and the Social Brain: Integrating Psychological and Evolutionary Perspectives. British Journal of Psychology 103(2), 68–149 (2012)
15. Wilson, C., Sala, A., Puttaswamy, K.P.N., Zhao, B.Y.: Beyond Social Graphs: User Interactions in Online Social Networks and Their Implications. ACM Transactions on the Web 6(4), 1–31 (2012)